Monotone Circuit Lower Bounds from Robust Sunflowers

Bruno Pasqualotto Cavalar*, Mrinal Kumar**, and Benjamin Rossman***

Abstract. Robust sunflowers are a generalization of combinatorial sunflowers that have applications in monotone circuit complexity [19], DNF sparsification [8], randomness extractors [12], and recent advances on the Erdős-Rado sunflower conjecture [3,13,16]. The recent breakthrough of Alweiss, Lovett, Wu and Zhang [3] gives an improved bound on the maximum size of a w-set system that excludes a robust sunflower. In this paper, we use this result to obtain an $\exp(n^{1/2-o(1)})$ lower bound on the monotone circuit size of an explicit n-variate monotone function, improving the previous record $\exp(n^{1/3-o(1)})$ of Harnik and Raz [9]. We also show an $\exp(\Omega(n))$ lower bound on the monotone *arithmetic* circuit size of a related polynomial. Finally, we introduce a notion of robust clique-sunflowers and use this to prove an $n^{\Omega(k)}$ lower bound on the monotone circuit size of the CLIQUE function for all $k \leqslant n^{1/3-o(1)}$, strengthening the bound of Alon and Boppana [1].

1 Introduction

A monotone Boolean circuit is a Boolean circuit with AND and OR gates but no negations (NOT gates). Although a restricted model of computation, monotone Boolean circuits seem a very natural model to work with when computing *monotone* Boolean functions, i.e., Boolean functions $f : \{0,1\}^n \to \{0,1\}$ such that for all pairs of inputs $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \{0,1\}^n$ where $a_i \leqslant b_i$ for every i, we have $f(a_1, a_2, \ldots, a_n) \leqslant f(b_1, b_2, \ldots, b_n)$. Many natural and well-studied Boolean functions such as Clique and Majority are monotone.

Monotone Boolean circuits have been very well studied in Computational Complexity over the years, and continue to be one of the few seemingly largest natural sub-classes of Boolean circuits for which we have exponential lower bounds. This line of work started with a very influential paper of Razborov [18] who proved a super-polynomial $n^{\Omega(\log n)}$ lower bound on the size of monotone circuits computing the $\mathsf{Clique}_{k,n}$ function for $k \leqslant \log n$. Prior to Razborov's result, we didn't even have super-linear lower bounds for monotone circuits with the best bound being a lower bound of $4n$ due to Tiekenheinrich [22]. Further progress in this line of work included the results of Andreev [4] who proved an exponential lower bound for another explicit function. Alon and Boppana [1] extended Razborov's result by proving an $n^{\Omega(\sqrt{k})}$ lower bound for $\mathsf{Clique}_{k,n}$ for all $k \leqslant n^{2/3-o(1)}$. These state of art for monotone circuit lower bounds saw a further quantitative improvement in a work of Harnik and Raz [9] who proved a lower bound of $2^{\Omega((n/\log n)^{1/3})}$ for an explicit n-variate function defined using a small probability space of random variables with bounded independence. However, to this day, the question of proving truly exponential lower bounds for monotone circuits (of the form $2^{\Omega(n)}$) for an explicit n-variate function) remains open! (Truly exponential lower bounds for monotone *formulas* were obtained only recently [15].)

In the present paper, we are able to improve the best known by proving the first $2^{\Omega(n^{1/2}/(\log n)^{3/2})}$ lower bound for an explicit monotone Boolean function (Section 2). The function is based on the same construction first considered by Harnik and Raz [9], but our argument employs the approximation method of Razborov with recent improvements on robust sunflower bounds [3,16]. By applying the same technique with a variant of robust sunflowers that we call robust clique-sunflowers, we are able to prove an $n^{\Omega(k)}$ lower bound for

* University of São Paulo, Email: brunopc@ime.usp.br. This work was support by grants #2018/22257-7 and #2018/05557-7, Sao Paulo Research Foundation (FAPESP). The author also acknowledges CAPES (PROEX) for partial support of this work.

** IIT Bombay, Email: mrinalkumar08@gmail.com. A part of this work was done during a postdoctoral stay at University of Toronto.

*** University of Toronto, Email: ben.rossman@utoronto.ca. Supported by NSERC, Ontario Early Researcher Award and Sloan Research Fellowship.

the $\mathsf{Clique}_{k,n}$ function when $k \leqslant n^{1/3-o(1)}$, thus improving the result of Alan and Boppana when k is in this range (Appendix B). Finally, we are able to prove truly exponential lower bounds in the monotone arithmetic setting to a fairly general family of polynomials, which shares some similarities to the Harnik and Raz function (Section 3).

1.1 Monotone circuit lower bounds and sunflowers

The original lower bound for $\mathsf{Clique}_{k,n}$ due to Razborov employed a technique which came to be known as the *approximation method*. Given a monotone circuit C of "small size", it consists into constructing gate-by-gate, in a bottom-up fashion, another circuit $\widetilde{C}$ that approximates C on most inputs of interest. One then exploits the structure of this *approximator circuit* to prove that it differs from $\mathsf{Clique}_{k,n}$ on most inputs of interest, thus implying that no "small" circuit can compute this function. This technique was leveraged to obtain lower bounds for a host of other monotone problems [1]. See Section 2 for a more detailed description of the method.

A crucial step in Razborov's proof involved the sunflower lemma due to Erdös and Rado. A family $\mathcal{F}$ of subsets of $[n]$ *sunflower* if there exists a set Y such that $F_1 \cap F_2 = Y$ for every $F_1, F_2 \in \mathcal{F}$. The sets of $\mathcal{F}$ are called *petals* and the set $Y = \bigcap \mathcal{F}$ is called the *core*.

Theorem 1 (Erdös and Rado [6]). *Let $\mathcal{F}$ be a family of subsets of $[n]$, each of cardinality at most ℓ. If $|\mathcal{F}| \geqslant \ell!(r-1)^\ell$, then $\mathcal{F}$ contains a sunflower of r petals.*

Informally, the sunflower lemma allows one to prove that a monotone function can be approximated by one with fewer minterms by means of the "plucking" procedure: if the function has too many (more than $\ell!(r-1)^\ell$) minterms of size ℓ, then it contains a sunflower with r petals; remove all the petals, replacing them with the core. One can then prove that this procedure does not introduce many errors.

The notion of *robust sunflowers* was introduced by the second author in [19], to achieve better bounds via the approximation method on the monotone circuit size of $\mathsf{Clique}_{k,n}$ when the negative instances are an Erdős-Rényi graphs $\boldsymbol{G}_{n,p}$ below the k-clique threshold.[1] A family $\mathcal{F} \subseteq 2^{[n]}$ is called a (p,ε)*-robust sunflower* if

$$\mathop{\mathbb{P}}_{\boldsymbol{W} \subseteq_p V} [\exists F \in \mathcal{F} : F \subseteq \boldsymbol{W} \cup Y] \geqslant 1 - \varepsilon,$$

where $Y := \bigcap \mathcal{F}$ and $\boldsymbol{W}$ is a p-random subset of $[n]$. Henceforth, we consistently write random objects using boldface symbols (such as $\boldsymbol{W}$, $\boldsymbol{G}_{n,p}$, etc). A corresponding bound for the appearance of robust sunflowers in large families was also proved in [19].

Theorem 2 ([19]). *Let $\mathcal{F} \subseteq \binom{[n]}{\ell}$ be such that $|\mathcal{F}| \geqslant \ell!(2\log(1/\varepsilon)/p)^\ell$. Then $\mathcal{F}$ contains a (p,ε)-robust sunflower.*

For many choice of parameters p and ε, this bound is better than the one by Erdös and Rado, thus leading to better approximation bounds. In a recent breakthrough, this result was significantly improved by Alweiss, Lovett, Wu and Zhang [3].

Theorem 3 ([3]). *Let $\mathcal{F} \subseteq \binom{[n]}{\ell}$ be such that $|\mathcal{F}| \geqslant (\log \ell)^\ell \cdot (\log\log \ell \cdot \log(1/\varepsilon)/p)^{O(\ell)}$. Then $\mathcal{F}$ contains a (p,ε)-robust sunflower.*

Soon afterwards, Rao [16] provided an alternative proof which slightly improved the bound. It is this bound we are going to use, which we introduce in the next section.[2]

[1] Robust sunflowers were called *quasi-sunflowers* in [19,8,12,13] and *approximate sunflowers* in [14]. Following Alweiss *et al* [3], we adopt the new name *robust sunflower*.

[2] Crucially for our application, the $O(\ell)$ exponent in the bound of Theorem 3 is only 2ℓ when $\varepsilon = 2^{-\Omega(\ell)}$. To get any improvement over the Harnik-Raz bound, we require $\ell + o(\ell)$, which is given by the result of Rao [16].

1.2 Slice sunflowers

In what follows, let m be a positive integer such that $m < n$.

Definition 1. *Let $\mathcal{F}$ be a family of subsets of $[n]$ and let $Y := \bigcap \mathcal{F}$. Let also $\boldsymbol{W} \subseteq [n]$ be a set of size m chosen uniformly at random. The family $\mathcal{F}$ is called a* (m, ε)-slice-sunflower *if*

$$\mathbb{P}_{\boldsymbol{W}} \left[\exists F \in \mathcal{F} : F \subseteq \boldsymbol{W} \cup Y \right] \geqslant 1 - \varepsilon.$$

Theorem 4 ([16]). *There exists an universal constant $B > 0$ such that the following holds. Let $p \in (0, 1)$ and let $\mathcal{F} \subseteq \binom{[n]}{\ell}$ be such that $|\mathcal{F}| \geqslant (Bx \log x)^{\ell}$, where $x = \log(\ell/\varepsilon)/p$. Then $\mathcal{F}$ contains a (m, ε)-slice-sunflower, where $m = \lfloor np \rfloor$.*

The theorem above is implicit in Rao [16]. For this reason, we include most of its proof in Appendix A, closely following the argument and notation of [16].

2 Harnik-Raz function

The strongest lower bound known for monotone circuits computing an explicit n-variate monotone Boolean function is $\exp\left(\Omega\left((n/\log n)^{1/3}\right)\right)$, and was obtained by Harnik and Raz [9]. In this section, we prove (Theorem 5) a lower bound of $\exp(\Omega(n^{1/2}/(\log n)^{3/2}))$ for the same Boolean function that Harnik and Raz considered. We apply the *method of approximations* of Razborov [18] and the new *robust sunflower* bound (Theorem 4) due to Alweiss, Lovett, Wu and Zhang [3] and Rao [16]. We do not expect that a lower bound better than $\exp(n^{1/2-o(1)})$ can be obtained by this technique, even with better sunflower bounds.

We start by giving a high level outline of the proof. We define the Harnik-Raz function $f_{\mathrm{HR}} : \{0,1\}^n \to \{0,1\}$ and find two distributions $\boldsymbol{Y}$ and $\boldsymbol{N}$ with support in $\{0,1\}^n$ satisfying the following properties:

- f_{HR} outputs 1 on $\boldsymbol{Y}$ with high probability (Lemma 1);
- f_{HR} outputs 0 on $\boldsymbol{N}$ with high probability (Lemma 2).

Because of these properties, the distribution $\boldsymbol{Y}$ is called the *positive test distribution*, and $\boldsymbol{N}$ is called the *negative test distribution*. We also define a set of monotone Boolean functions called *approximators*, and we show that:

- every approximator commits many mistakes on either $\boldsymbol{Y}$ or $\boldsymbol{N}$ with high probability (Lemma 8);
- every Boolean function computed by a "small" monotone circuit agrees with an approximator on both $\boldsymbol{Y}$ and $\boldsymbol{N}$ with high probability (Lemma 9).

Together these suffice for proving that "small" circuits cannot compute f_{HR}. The crucial part where the robust sunflower result comes into play is in the second item.

2.1 Technical preliminaries

For $A \subseteq [n]$, let $x_A \in \{0,1\}^n$ be the binary vector with support in A. For a set $A \in 2^{[n]}$, let $\lceil A \rceil$ be the indicator function satisfying

$$\lceil A \rceil(x) = 1 \iff x_A \leqslant x.$$

Define also $\{0,1\}^n_{=m} := \left\{ x_A : A \in \binom{n}{m} \right\}$. For a monotone Boolean function $f : \{0,1\}^n \to \{0,1\}$, let $\mathcal{M}(f)$ denote the set of minterms of f, and let $\mathcal{M}_\ell(f) := \mathcal{M}(f) \cap \{0,1\}^n_{=\ell}$. Elements of $\mathcal{M}_\ell(f)$ are called ℓ-minterms of f. In what follows, we will mostly ignore ceilings and floors for the sake of convenience, since these do not make any substantial difference in the final calculations.

2.2 The function

We start by describing the construction of the function $f_{\rm HR} : \{0,1\}^n \to \{0,1\}$ considered by Harnik and Raz [9]. In what follows, we suppose n is a prime power and let $\mathbb{F}_n$ be the field of n elements. Moreover, we fix two positive integers c and k, of which values the construction of $f_{\rm HR}$ depends. Let $S_{\rm HR}$ be the set of polynomials $P \in \mathbb{F}_n[x]$ with degree at most $c-1$ such that $|\{P(1), P(2), \ldots, P(k)\}| \geqslant k/2$. The function $f_{\rm HR}$ is defined as follows:[3]

$$f_{\rm HR}(x_1, \ldots, x_n) := \bigvee_{P \in S_{\rm HR}} \bigwedge_{j=1}^{k} x_{P(j)}.$$

In other words, given $x \in \{0,1\}^n$, we have $f_{\rm HR}(x) = 1$ if and only if there exists $P \in S_{\rm HR}$ such that $x_{P(i)} = 1$ for all $i \in [k]$. Note that, since $|S_{\rm HR}| \leqslant n^c$, the function $f_{\rm HR}$ has at most n^c minterms. Moreover, every minterm of $f_{\rm HR}$ has Hamming weight at least $k/2$.

We now define the positive and negative test distributions. Let $\boldsymbol{Y} \in \{0,1\}^n$ be the random variable which chooses a polynomial $\boldsymbol{P} \in \mathbb{F}_n[x]$ with degree at most $c-1$ uniformly at random, and maps it into the binary input $x_{\{\boldsymbol{P}(1), \boldsymbol{P}(2), \ldots, \boldsymbol{P}(k)\}} \in \{0,1\}^n$. It is known [2] that the random variables $\boldsymbol{P}(1), \boldsymbol{P}(2), \ldots, \boldsymbol{P}(k) \in [n]$ are c-wise independent. Moreover, each $\boldsymbol{P}(i)$ is uniform in $[n]$. Let

$$p := n^{-4c/k} \quad \text{and} \quad m := \lfloor np \rfloor.$$

Let also $\boldsymbol{N}$ be the distribution which chooses an input from $\{0,1\}^n_{=m}$ uniformly at random. For a Boolean function f and a probability distribution $\boldsymbol{\mu}$ on the inputs on f, we write $f(\boldsymbol{\mu})$ to denote the random variable which evaluates f on a random instance of $\boldsymbol{\mu}$. Harnik and Raz proved that $f_{\rm HR}$ outputs 1 on $\boldsymbol{Y}$ with high probability.

Lemma 1 (Claim 4.2 in [9]). *We have* $\mathbb{P}[f_{\rm HR}(\boldsymbol{Y}) = 1] \geqslant 1 - k/n$.

We now claim that $f_{\rm HR}$ also outputs 0 on $\boldsymbol{N}$ with high probability.

Lemma 2. *We have* $\mathbb{P}[f_{\rm HR}(\boldsymbol{N}) = 0] \geqslant 1 - n^{-c}$.

Proof. Let x_A be an input sampled from $\boldsymbol{N}$. Observe that $f_{\rm HR}(x_A) = 1$ only if there exists a minterm x of $f_{\rm HR}$ such that $x \leqslant x_A$. Since all minterms of $f_{\rm HR}$ have Hamming weight at least $k/2$ and $f_{\rm HR}$ has at most n^c minterms, we have

$$\mathbb{P}[f_{\rm HR}(\boldsymbol{N}) = 1] \leqslant n^c \cdot \frac{\binom{n-k/2}{m-k/2}}{\binom{n}{m}} \leqslant n^c \cdot \left(\frac{m}{n}\right)^{k/2} \leqslant n^{-c}.$$

As a consequence of Lemmas 1 and 2, we obtain the following result.

Lemma 3. *For n sufficiently large, we have* $\mathbb{P}[f_{\rm HR}(\boldsymbol{Y}) = 1] + \mathbb{P}[f_{\rm HR}(\boldsymbol{N}) = 0] \geqslant 9/5$.

2.3 A closure operator

In this section, we describe a closure operator in the lattice of monotone Boolean functions. We prove that the closure $\mathrm{cl}(f)$ of a monotone Boolean function f is a good approximation for f on the negative test distribution (Lemma 4), and we give a bound on the size of the set of minterms of *closed* monotone functions. This bound makes use of the robust sunflower lemma of Alweiss, Lovett, Wu and Zhang [3] and Rao [16] (Theorem 4), and is crucial to bounding errors of approximation (Lemma 7). Throughout this section, we let

$$\varepsilon := n^{-3c}.$$

[3] As remarked in Harnik and Raz [9], we could replace the set of polynomials of degree at most $c-1$ by any probability space of k random variables which are c-wise independent.

Definition 2. *We say that a monotone function* $f : \{0,1\}^n \to \{0,1\}$ *is* ε-closed *if, for every* $A \in \binom{[n]}{\leqslant c}$*, we have*

$$\mathbb{P}[\ f(\boldsymbol{N} \vee x_A) = 1\ ] \geqslant 1 - \varepsilon \implies f(x_A) = 1.$$

This means that for, an ε-closed function, we always have $\mathbb{P}[f(\boldsymbol{N} \vee x_A) = 1] \notin [1-\varepsilon, 1)$ when $|A| \leqslant c$. Note morever that if f, g are both ε-closed monotone Boolean functions, then so is $f \wedge g$. Therefore, there exists a unique minimum closed function $\mathrm{cl}(f)$ satisfying $f \leqslant \mathrm{cl}(f)$. We call $\mathrm{cl}(f)$ the *closure* of f. We now give a bound on the error of approximating f by $\mathrm{cl}(f)$ under the distribution $\boldsymbol{N}$.

Lemma 4. *For every monotone* $f : \{0,1\}^n \to \{0,1\}$*, we have*

$$\mathbb{P}\left[f(\boldsymbol{N}) = 0 \text{ and } \mathrm{cl}(f)(\boldsymbol{N}) = 1\right] \leqslant n^{-2c}.$$

Proof. We first prove that there exists a positive integer t and sets $A_1, \dots, A_t$ and monotone functions $h_0, h_1, \dots, h_t : \{0,1\}^n \to \{0,1\}$ such that

1. $h_0 = f$,
2. $h_i = h_{i-1} \vee \lceil A_i \rceil$,
3. $\mathbb{P}[h_{i-1}(\boldsymbol{N} \cup x_{A_i}) = 1] \geqslant 1 - \varepsilon$,
4. $h_t = \mathrm{cl}(f)$.

Indeed, if h_{i-1} is not closed, there exists $A_i \in \binom{[n]}{\leqslant c}$ such that $\mathbb{P}[h_{i-1}(\boldsymbol{N} \cup x_{A_i}) = 1] \geqslant 1-\varepsilon$ but $h_{i-1}(x_{A_i}) = 0$. We let $h_i := h_{i-1} \vee \lceil A_i \rceil$. Clearly, we have that h_t is closed, and that the value of t is at most the number of subsets of $[n]$ of size at most c. Therefore, we get $t \leqslant \sum_{j=0}^{c} \binom{n}{j}$. Moreover, by induction we obtain that $h_i \leqslant \mathrm{cl}(f)$ for every $i \in [t]$. It follows that $h_t = \mathrm{cl}(f)$. Now, observe that

$$\begin{aligned}
\mathbb{P}\left[f(\boldsymbol{N}) = 0 \text{ and } \mathrm{cl}(f)(\boldsymbol{N}) = 1\right] &\leqslant \sum_{i=1}^{t} \mathbb{P}\left[f_{i-1}(\boldsymbol{N}) = 0 \text{ and } f_i(\boldsymbol{N}) = 1\right] \\
&= \sum_{i=1}^{t} \mathbb{P}\left[f_{i-1}(\boldsymbol{N}) = 0 \text{ and } x_{A_i} \subseteq \boldsymbol{N}\right] \\
&\leqslant \sum_{i=1}^{t} \mathbb{P}\left[f_{i-1}(\boldsymbol{N} \cup x_{A_i}) = 0\right] \\
&\leqslant \varepsilon \sum_{j=0}^{c} \binom{n}{j} \leqslant n^{-2c}.
\end{aligned}$$

We now give a bound on the size of the set of ℓ-minterms of an ε-closed function. This bound is dependent on the robust sunflower theorem (Theorem 4) and it is the key property that allows us to obtain the better lower bound againt f_{HR} (Theorem 5).

Lemma 5. *Let* $B > 0$ *be as in Theorem 4. If a monotone function* $f : \{0,1\}^n \to \{0,1\}$ *is* ε*-closed, then, for all* $\ell \in [c]$*, we have*

$$|\mathcal{M}_\ell(f)| \leqslant \left(B \frac{\log(\ell/\varepsilon)}{p} \log\left(\frac{\log(\ell/\varepsilon)}{p}\right)\right)^{\ell}.$$

Proof. Fix $\ell \in [c]$. Suppose we have $|\mathcal{M}_\ell(f)| > (C\log(\ell/\varepsilon)/p \log(\log(\ell/\varepsilon)/p))^\ell$. Consider also the family $\mathcal{F} := \left\{A \in \binom{[n]}{\ell} : x_A \in \mathcal{M}_\ell(f)\right\}$. Observe that $|\mathcal{F}| = |\mathcal{M}_\ell(f)|$. By Theorem 4, there exists a (m, ε)-slice-sunflower $\mathcal{F}' \subseteq \mathcal{F}$. Let $Y := \bigcap \mathcal{F}'$ and let $W \in \binom{[n]}{m}$ be chosen uniformly at random. We have

$$\begin{aligned}
\mathbb{P}[f(\boldsymbol{N} \vee x_Y) = 1] &\geqslant \mathbb{P}[\exists x \in \mathcal{M}_\ell(f) : x \leqslant \boldsymbol{N} \vee x_Y] \\
&= \mathbb{P}[\exists F \in \mathcal{F} : F \subseteq W \cup Y] \\
&\geqslant \mathbb{P}[\exists F \in \mathcal{F}' : F \subseteq W \cup Y] \\
&\geqslant 1 - \varepsilon.
\end{aligned}$$

Therefore, since f is ε-closed, we get that $f(x_Y) = 1$. However, since $Y = \bigcap \mathcal{F}'$, there exists $F \in \mathcal{F}'$ such that $Y \subsetneq F$. This is a contradiction, because x_F is a minterm of f.

2.4 Trimmed monotone functions

In this section, we define a *trimming* operation for Boolean functions. We will bound the probability that a *trimmed* function gives the correct output on the distribution $\boldsymbol{Y}$, and we will give a bound on the error of approximating a Boolean function f by the trimming of f on that same distribution.

Definition 3. *We say that a monotone function $f \in \{0,1\}^n \to \{0,1\}$ is* trimmed *if all the minterms of f have size at most $c/2$. We define the trimming operation* $\mathrm{trim}(f)$ *as follows:*

$$\mathrm{trim}(f) := \bigvee_{\ell=1}^{c/2} \bigvee_{A \in \mathcal{M}_\ell(f)} \lceil A \rceil.$$

That is, the trim operation takes out from f all the minterms of size larger than $c/2$, yielding a trimmed function. We will first prove the following claim.

Claim. For every monotone function $f : \{0,1\}^n \to \{0,1\}$ and $\ell \leqslant c$, we have

$$\mathbb{P}[\exists x \in \mathcal{M}_\ell(f) : x \leqslant \boldsymbol{Y}] \leqslant \left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(f)|\,.$$

Proof. Recall (Section 2.2) that the distribution $\boldsymbol{Y}$ takes a polynomial $\boldsymbol{P} \in \mathbb{F}_n[x]$ with degree at most $c-1$ uniformly at random and returns the binary vector $x_{\{\boldsymbol{P}(1),\boldsymbol{P}(2),\ldots,\boldsymbol{P}(k)\}} \in \{0,1\}^n$. Let $A \in \binom{[n]}{\ell}$ for $\ell \leqslant c$. Observe that $x_A \leqslant \boldsymbol{Y}$ if and only if $A \subseteq \{\boldsymbol{P}(1), \boldsymbol{P}(2), \ldots, \boldsymbol{P}(k)\}$. Therefore, if $x_A \leqslant \boldsymbol{Y}$, then there exists indices $\{j_1, \ldots, j_\ell\}$ such that $\{\boldsymbol{P}(j_1), \boldsymbol{P}(j_2), \ldots, \boldsymbol{P}(j_\ell)\} = A$. Since $\ell \leqslant c$, we get by the c-wise independence of $\boldsymbol{P}(1), \ldots, \boldsymbol{P}(k)$ that the random variables $\boldsymbol{P}(j_1), \boldsymbol{P}(j_2), \ldots, \boldsymbol{P}(j_\ell)$ are independent. It follows that

$$\mathbb{P}[\{\boldsymbol{P}(j_1), \boldsymbol{P}(j_2), \ldots, \boldsymbol{P}(j_\ell)\} = A] = \frac{\ell!}{n^\ell}.$$

Therefore, we have

$$\mathbb{P}[x_A \leqslant \boldsymbol{Y}] = \mathbb{P}[A \subseteq \{\boldsymbol{P}(1), \boldsymbol{P}(2), \ldots, \boldsymbol{P}(k)\}] \leqslant \binom{k}{\ell}\frac{\ell!}{n^\ell} \leqslant \left(\frac{k}{n}\right)^\ell.$$

We thus obtain

$$\mathbb{P}[\exists x \in \mathcal{M}_\ell(f) : x \leqslant \boldsymbol{Y}] \leqslant \left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(f)|\,.$$

We are now able to bound the probability that a trimmed Boolean function gives the correct output on distribution $\boldsymbol{Y}$ and give a bound on the approximation error of the trimming operation.

Lemma 6. *If a monotone function $f \in \{0,1\}^n \to \{0,1\}$ is trimmed and $f \neq \mathbb{1}$ (i.e., f is not identically 1), then*

$$\mathbb{P}\left[f(\boldsymbol{Y}) = 1\right] \leqslant \sum_{\ell=1}^{c/2} \left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(f)|\,.$$

Proof. It suffices to see that, since f is trimmed, if $f(\boldsymbol{Y}) = 1$ and $f \neq \mathbb{1}$ then there exists a minterm x of f with Hamming weight between 1 and $c/2$ such that $x \leqslant \boldsymbol{Y}$. The result follows by the claim above.

Lemma 7. *Let $f \in \{0,1\}^n \to \{0,1\}$ be a monotone function, all of whose minterms have Hamming weight at most c. We have*

$$\mathbb{P}\left[f(\boldsymbol{Y}) = 1 \text{ and } \operatorname{trim}(f)(\boldsymbol{Y}) = 0\right] \leqslant \sum_{\ell=c/2}^{c} \left(\frac{k}{n}\right)^{\ell} |\mathcal{M}_\ell(f)| .$$

Proof. If we have $f(\boldsymbol{Y}) = 1$ and $\operatorname{trim}(f)(\boldsymbol{Y}) = 0$, then there was a minterm x of f with Hamming weight larger than $c/2$ that was removed by the trimming process. Therefore, since $|x| \leqslant c$ by assumption, the result follows by the claim above.

2.5 The approximators

Let $\mathcal{A} := \{\operatorname{trim}(\operatorname{cl}(f)) : f : \{0,1\}^n \to \{0,1\} \text{ is monotone}\}$. Functions in $\mathcal{A}$ will be called *approximators*. We define the *approximating* operations $\sqcup, \sqcap : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ as follows: for $f, g \in \mathcal{A}$, let

$$f \sqcup g := \operatorname{trim}(\operatorname{cl}(f \vee g)),$$
$$f \sqcap g := \operatorname{trim}(\operatorname{cl}(f \wedge g)).$$

Observe that every input function x_i is an approximator. Therefore, we can replace each gate of a monotone $\{\vee, \wedge\}$-circuit C by its corresponding approximating gate, thus obtaining a $\{\sqcup, \sqcap\}$-circuit $C^{\mathcal{A}}$ which computes an approximator.

The rationale for choosing this set of approximators is as follows. By letting approximators be the trimming of a closed function, we are able to plug the bound on the set of ℓ-minterms given by the robust sunflower lemma (Lemma 5) on Lemmas 6 and 7, since the trimming operation can only *reduce* the set of minterms. Moreover, since trimmings can only help to get a negative answer on the negative test distribution, we can safely apply Lemma 4 when bounding the errors of approximation.

2.6 The lower bound

In this section, we will prove that the function f_{HR} requires monotone circuits of size $2^{\Omega(c)}$. This is nearly tight: since it has at most $2^{c \log n}$ minterms, it can be computed by a DNF of this size. By properly choosing c and k, this will imply the promised $\exp(\Omega(n^{1/2-o(1)}))$ lower bound for the Harnik-Raz function. Our strategy will be to bound the error of approximating a monotone circuit C by $C^{\mathcal{A}}$ (Lemma 9) and to show that approximators err on the input distributions with non-negligible probability (Lemma 8).

First, we fix some parameters. Choose B as in Lemma 5. We also let

$$c := \frac{1}{6Be^{1/B}} \left(\frac{n}{(\log n)^3}\right)^{1/2}, \qquad k := \left(\frac{n}{\log n}\right)^{1/2}.$$

For simplicity, we assume these values are integers. We clearly have $c < k$. Moreover, observe that, because of this choice of parameters, we have $p = \Omega(1)$. Indeed, we have

$$p = n^{-4c/k} = n^{-2/(3Be^{1/B} \log n)} = e^{-2/(3Be^{1/B})} \geqslant e^{-1/B}.$$

What we now want to do is to show that, when f is an approximator, the bound of Lemma 6 can be replaced by $1/2$, and also that, when f is an ε-closed function, the bound of Lemma 7 can be replaced by $2^{-\Omega(c)}$. We will first need to bound the sequence s_ℓ, defined as follows. For every $1 \leqslant \ell \leqslant c$, let

$$s_\ell := \left(\frac{k}{n}\right)^{\ell} \cdot \left(B \frac{\log(c/\varepsilon)}{p} \log\left(\frac{\log(c/\varepsilon)}{p}\right)\right)^{\ell}.$$

Note that, when $f : \{0,1\}^n \to \{0,1\}$ is an ε-closed monotone function, we get by Lemma 5 that $\left(\frac{k}{n}\right)^{\ell} |\mathcal{M}_\ell(f)| \leqslant s_\ell$. In other words, the summands of Lemma 6 and Lemma 7 can be replaced by s_ℓ in some applications.

Observe moreover that $s_\ell = (s_1)^\ell$. Now we are going to show that, for n sufficiently large, we have $s_1 \leqslant 1/3$, which implies $s_\ell \leqslant 3^{-\ell}$. First, observe that

$$\log(c/\varepsilon)/p = \log(n^{3c}c)/p \leqslant \log(n^{4c})/p = \frac{4c}{p}\log n.$$

Moreover, we have

$$\log\left(\log(c/\varepsilon)/p\right) = \log\left(\frac{4c}{p}\log n\right) = \frac{1}{2}\log n - \frac{1}{2}\log\log n + O(1) \leqslant \frac{1}{2}\log n,$$

for n sufficiently large. From the previous two inequalities, we obtain for n sufficiently large that

$$s_1 = B\cdot\frac{k}{n}\cdot\frac{\log(c/\varepsilon)}{p}\log\left(\frac{\log(c/\varepsilon)}{p}\right) \leqslant \frac{2B}{p}\cdot\frac{ck(\log n)^2}{n} \leqslant 1/3,$$

as desired. We are now able to give the desired bounds for the accuracy of the approximators and for the approximation errors.

Lemma 8 (Approximators make many errors). *For every approximator $f \in \mathcal{A}$, we have*

$$\mathbb{P}[f(\boldsymbol{Y}) = 1] + \mathbb{P}[f(\boldsymbol{N}) = 0] \leqslant 3/2.$$

Proof. Let $f \in \mathcal{A}$. By definition, there exists an ε-closed function h such that $f = \mathrm{trim}(h)$. Observe that $\mathcal{M}_\ell(f) \subseteq \mathcal{M}_\ell(h)$ for every $\ell \in [c]$. Hence, applying Lemma 6 and the bounds for s_ℓ, we obtain that, if $f \neq \mathbb{1}$, we have

$$\mathbb{P}[f(\boldsymbol{Y}) = 1] \leqslant \sum_{\ell=1}^{c/2}\left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(h)| \leqslant \sum_{\ell=1}^{c/2} s_\ell \leqslant \sum_{\ell=1}^{c/2} 3^{-\ell} \leqslant 1/2.$$

Therefore, for every $f \in \mathcal{A}$ we have $\mathbb{P}[f(\boldsymbol{Y}) = 1] + \mathbb{P}[f(\boldsymbol{N}) = 0] \leqslant 1 + 1/2 \leqslant 3/2$.

Lemma 9 (C is well-approximated by $C^{\mathcal{A}}$). *Let C be a monotone circuit. We have*

$$\mathbb{P}[C(\boldsymbol{Y}) = 1 \text{ and } C^{\mathcal{A}}(\boldsymbol{Y}) = 0] + \mathbb{P}[C(\boldsymbol{N}) = 0 \text{ and } C^{\mathcal{A}}(\boldsymbol{N}) = 1] \leqslant \mathrm{size}(C)\cdot 2^{-\Omega(c)}.$$

Proof. We begin by bounding the approximation errors under the distribution $\boldsymbol{Y}$. We will show that, for two approximators $f, g \in \mathcal{A}$, if $f \vee g$ accepts an input from $\boldsymbol{Y}$, then $f \sqcup g$ rejects that input with probability at most $2^{-\Omega(c)}$, and that the same holds for the approximation $f \sqcap g$.

First note that, if $f, g \in \mathcal{A}$, then all the minterms of both $f \vee g$ and $f \wedge g$ have Hamming weight at most c, since f and g are trimmed. Let now $h = \mathrm{cl}(f \vee g)$. Observe that $h \geqslant f \vee g$ and that $\mathrm{trim}(h) = f \sqcup g$. Therefore, if $(f \sqcup g)(x) < (f \vee g)(x)$ then $\mathrm{trim}(h)(x) < h(x)$. Observing that h is closed, we obtain the following inequality by Lemma 7 and the bounds on s_ℓ:

$$\mathbb{P}\left[(f\vee g)(\boldsymbol{Y}) = 1 \text{ and } (f \sqcup g)(\boldsymbol{Y}) = 0\right] \leqslant \sum_{\ell=c/2}^{c}\left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(h)| \leqslant \sum_{\ell=c/2}^{c} s_\ell \leqslant \sum_{\ell=c/2}^{c} 3^{-\ell} = 2^{-\Omega(c)}.$$

The same argument shows $\mathbb{P}\left[(f \wedge g)(\boldsymbol{Y}) = 1 \text{ and } (f \sqcap g)(\boldsymbol{Y}) = 0\right] = 2^{-\Omega(c)}$. Since there are $\mathrm{size}(C)$ gates in C, this implies that

$$\mathbb{P}[C(\boldsymbol{Y}) = 1 \text{ and } C^{\mathcal{A}}(\boldsymbol{Y}) = 0] \leqslant \mathrm{size}(C)\cdot 2^{-\Omega(c)}.$$

Similarly, to bound the approximation errors under $\boldsymbol{N}$, note that $(f \vee g)(x) = 0$ and $(f \sqcup g)(x) = 1$ only if $\mathrm{cl}(f \vee g)(x) \neq (f \vee g)(x)$, since trimming a Boolean function can only increase the probability that it rejects an input. Therefore, by Lemma 4 we obtain

$$\mathbb{P}\left[(f \vee g)(\boldsymbol{N}) = 0 \text{ and } (f \sqcup g)(\boldsymbol{N}) = 1\right] \leqslant n^{-2c} \leqslant 2^{-\Omega(c)}.$$

Once again, doing this approximation for every gate in C implies

$$\mathbb{P}[C(\boldsymbol{N}) = 0 \text{ and } C^{\mathcal{A}}(\boldsymbol{N}) = 1] \leqslant \mathrm{size}(C)\cdot 2^{-\Omega(c)}.$$

This finishes the proof.

We can now finally obtain the desired lower bound against f_{HR}.

Theorem 5. *Any monotone circuit computing f_{HR} has size $2^{\Omega(c)} = 2^{\Omega(n^{1/2}/(\log n)^3)}$.*

Proof. Let C be a monotone circuit computing f_{HR}. For n sufficiently large, we have

$$\begin{aligned}9/5 &\leqslant \mathbb{P}[f_{\mathrm{HR}}(\boldsymbol{Y}) = 1] + \mathbb{P}[f_{\mathrm{HR}}(\boldsymbol{N}) = 0] \\ &\leqslant \mathbb{P}[C(\boldsymbol{Y}) = 1 \text{ and } C^{\mathcal{A}}(\boldsymbol{Y}) = 0] + \mathbb{P}[C^{\mathcal{A}}(\boldsymbol{Y}) = 1] \\ &+ \mathbb{P}[C(\boldsymbol{N}) = 0 \text{ and } C^{\mathcal{A}}(\boldsymbol{N}) = 1] + \mathbb{P}[C^{\mathcal{A}}(\boldsymbol{N}) = 0] \\ &= 3/2 + \mathrm{size}(C) 2^{-\Omega(c)}.\end{aligned}$$

This implies $\mathrm{size}(C) = 2^{\Omega(c)}$.

2.7 Discussion

In this application, we chose the values of c and k to be roughly $\sqrt{n}$. We expect that, if c were chosen to be closer to n, the implied Harnik-Raz function would still have $2^{\Omega(c)}$ complexity, and thus one would be able to improve our bound. However, we do not think that the present technique would work for any $c > \sqrt{n}$, as it seems to require that $ck \leqslant n$. Therefore, in order to obtain a stronger bound to the Harnik-Raz function, we think one would need to consider a different technique.

3 Monotone arithmetic circuits

In this section, we give a short and simple proof of a truly exponential ($\exp(\Omega(n))$) lower bound for real monotone algebraic circuits computing a multilinear n variate polynomial. As we shall see, the lower bound argument holds for a general family of multilinear polynomials constructed in a very natural way from error correcting codes, and the similarities to the hard function used by Harnik and Raz in the Boolean setting is quite evident (see Section 2.2). In particular, our lower bound just depends on the rate and relative distance of the underlying code. We note that exponential lower bounds for monotone algebraic circuits are not new, and have been known since the 80's with various quantitative bounds. More precisely, Jerrum and Snir proved an $\exp(\Omega(\sqrt{n}))$ lower bound for an n variate polynomial in [11]. This bound was subsequently improved to a lower bound of $\exp(\Omega(n))$ by Raz and Yehudayoff in [17], via an extremely clever argument, which relied on deep and beautiful results on character sums over finite fields. A similar lower bound of $\exp(\Omega(n))$ was shown by Sirnivasan [20] using more elementary techniques building on a work of Yehudayoff [23].

We show a similar lower bound of $\exp(\Omega(n))$ via a simple and short argument, which holds in a somewhat general setting. Our contribution is just the simplicity, the (lack of) length of the argument and the observation that it holds for families of polynomials that can be constructed from any sufficiently *good* error correcting codes.

Definition 4 (From sets of vectors to polynomials). *Let $C \subseteq \mathbb{F}_q^n$ be an arbitrary subset of $\mathbb{F}_q^n$. Then, the polynomial P_C is a multilinear homogeneous polynomial of degree n on qn variables $\{x_{i,j} : i \in [q], j \in [n]\}$ and is defined as follows:*

$$P_C = \sum_{c \in C} \prod_{j \in [n]} x_{j,c(j)} \, .$$

Here, $c(j)$ is the j^{th} coordinate of c which is an element of $\mathbb{F}_q$, which we bijectively identify with the set $[q]$.

Here, we will be interested in the polynomial P_C when the set C is a *good* code, i.e it has high rate and high relative distance. The following observation summarizes the properties of P_C and relations between the properties of C and P_C.

Observation 6 (Codes vs Polynomials) *Let C be any subset of $\mathbb{F}_q^n$ and let P_C be the polynomial as defined in Definition 4. Then, the following statements are true:*

- *P_C is a multilinear homogeneous polynomial of degree equal to n with every coefficient being either 0 or 1.*
- *The number of monomials with non-zero coefficients in P_C is equal to the cardinality of C.*
- *If any two distinct vectors in C agree on at most k coordinates (i.e. C is a code of distance $n - k$), then the intersection of the support of any two monomials with non-zero coefficients in P_C has size at most k.*

The observation immediately follows from Definition 4. We note that we will work with monotone algebraic circuits here, and hence will interpret the polynomial P_C as a polynomial over the field of real numbers.

We now prove the following theorem, which essentially shows that for every code C with sufficiently good distance, any monotone algebraic circuit computing P_C must essentially compute it by computing each of its monomials separately, and taking their sum.

Theorem 7. *If any two distinct vectors in C agree on at most $n/3-1$ locations, then any monotone algebraic circuit for P_C has size at least $|C|$.*

The proof of this theorem crucially uses the following well known structural lemma about algebraic circuits. This lemma also plays a crucial role in the other proofs of exponential lower bounds for monotone algebraic circuits (e.g. [11,17,23,20]).

Lemma 10 (See Lemma 3.3 in [17]). *Let Q be a homogeneous multilinear polynomial polynomial of degree d computable by a homogeneous algebraic circuit of size s. Then, there are homogeneous polynomials $g_0, g_1, g_2, \ldots, g_s, h_0, h_1, h_2, \ldots, h_s$ of degree at least $d/3$ and at most $2d/3 - 1$ such that*

$$Q = \sum_{i=0}^{s} g_i \cdot h_i \,.$$

Moreover, if the circuit for Q is monotone, then each g_i and h_i is multilinear, variable disjoint and each one their non-zero coefficients is a positive real number.

We now use this lemma to prove Theorem 7.

Proof of Theorem 7. Let B be a monotone algebraic circuit for P_C of size s. We know from Observation 6 that P_C is a multilinear homogeneous polynomial of degree equal to n. This along with the monotonicity of B implies that B must be homogeneous and multilinear since there can be no cancellations in B. Thus, from (the moreover part of) Lemma 10 we know that P_C has a monotone decomposition of the form

$$P_C = \sum_{i=0}^{s} g_i \cdot h_i \,,$$

where, each g_i and h_i is multilinear, homogeneous with degree between $n/3$ and $2n/3 - 1$, g_i and h_i are variable disjoint. We now make the following claim.

Claim. Each g_i and h_i has at most one non-zero monomial.

We first observe that the claim immediately implies theorem 7: since every g_i and h_i has at most one non-zero monomial, their product $g_i h_i$ is just a monomial. Thus, the number of summands s needed in the decomposition above must be equal to the number of monomials in P_C, which is equal to $|C|$ from the second item in Observation 6.

We now prove the Claim.

Proof of Claim. The proof of the claim will be via contradiction. To this end, let us assume that there is an $i \in \{0, 1, 2, \ldots, s\}$ such that g_i has at least two distinct monomials with non-zero coefficients and let α and β be two of these monomials. Let γ be a monomial with non-zero coefficient in h_i . Since h_i is homogeneous with degree between $n/3$ and $2n/3 - 1$, we know that the degree of γ is at least $n/3$. Since we are in the monotone setting, we also know that each non-zero coefficient in any of the g_j and h_j is a positive real number. Thus, the monomials $\alpha \cdot \gamma$ and $\beta \cdot \gamma$ which have non-zero coefficients in the product $g_i \cdot h_i$ must have non-zero coefficient in P_C as well (since a monomial once computed cannot be cancelled out). But, the supports of $\alpha\gamma$ and $\beta\gamma$ overlap on γ which has degree at least $n/3$. This contradicts the fact that no two distinct monomials with non-zero coefficients in P_C share a sub-monomial of degree at least $n/3$ from the distance of C and the third item in Observation 6.

Theorem 7 when instantiated with an appropriate choice of the code C, immediately implies an exponential lower bound on the size of monotone algebraic circuits computing the polynomial P_C. Observe that the total number of variables in P_C is $N = qn$ and therefore, for the lower bound for P_C to be of the form $\exp(\Omega(N))$, we would require q, the underlying field size to be a constant. In other words, for any code of relative distance at least $2/3$ over a constant size alphabet which has exponentially many code words, we have a truly exponential lower bound.

The following theorem of Garcia and Stichtenoth [7] implies an explicit construction of such codes. The statement below is a restatement of their result by Cohen et al.[5].

Theorem 8 ([7] and [21]). *Let p be a prime number and let $m \in \mathbb{N}$ be even. Then, for every $0 < \rho < 1$ and a large enough integer n, there exists an explicit rate ρ linear error correcting block code $C : \mathbb{F}_{p^m}^n \to \mathbb{F}_{p^m}^{n/\rho}$ with distance*

$$\delta \geqslant 1 - \rho - \frac{1}{p^{m/2} - 1} \, .$$

The theorem has the following immediate corollary.

Corollary 1. *For every large enough constant q which is an even power of a prime, and for all large enough n, there exist explicit construction of codes $C \subseteq \mathbb{F}_q^n$ which have relative distance at least $2/3$ and $|C| \geqslant \exp(\Omega(n))$.*

By an explicit construction here, we mean that given a vector v of length n over $\mathbb{F}_q$, we can decide in deterministic polynomial time if $v \in C$. In the algebraic complexity literature, a polynomial P is said to be explicit, if given the exponent vector of a monomial, its coefficient in P can be computed in deterministic polynomial time. Thus, if a code C is explicit, then the corresponding polynomial P_C is also explicit in the sense described above. Therefore, we have the following corollary of Corollary 1 and Theorem 7.

Corollary 2. *There exists an explicit family $\{P_n\}$ of homogeneous multilinear polynomials such that for every large enough n, any monotone algebraic circuit computing the n variate polynomial P_n has size at least $\exp(\Omega(n))$.*

References

1. Alon, N., Boppana, R.B.: The monotone circuit complexity of Boolean functions. Combinatorica **7**(1), 1–22 (1987), https://doi.org/10.1007/BF02579196
2. Alon, N., Babai, L., Itai, A.: A fast and simple randomized parallel algorithm for the maximal independent set problem. J. Algorithms **7**(4), 567–583 (1986). https://doi.org/10.1016/0196-6774(86)90019-2, https://doi.org/10.1016/0196-6774(86)90019-2
3. Alweiss, R., Lovett, S., Wu, K., Zhang, J.: Improved bounds for the sunflower lemma. arXiv:1908.08483 (2019), https://arxiv.org/abs/1908.08483
4. Andreev, A.E.: A method for obtaining lower bounds on the complexity of individual monotone functions. Dokl. Akad. Nauk SSSR **282**(5), 1033–1037 (1985)

5. Cohen, G., Haeupler, B., Schulman, L.J.: Explicit binary tree codes with polylogarithmic size alphabet. In: Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing. pp. 535–544. STOC 2018, ACM. https://doi.org/10.1145/3188745.3188928, http://doi.acm.org/10.1145/3188745.3188928
6. Erdős, P., Rado, R.: Intersection theorems for systems of sets. J. London Math. Soc. **35**, 85–90 (1960), https://doi.org/10.1112/jlms/s1-35.1.85
7. Garcia, A., Stichtenoth, H.: A tower of artin-schreier extensions of function fields attaining the drinfeld-vladut bound. Inventiones Mathematicae **121**(1), 211–222 (1995)
8. Gopalan, P., Meka, R., Reingold, O.: DNF sparsification and a faster deterministic counting algorithm. Computational Complexity **22**(2), 275–310 (2013)
9. Harnik, D., Raz, R.: Higher lower bounds on monotone size. In: Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing. pp. 378–387. ACM, New York (2000). https://doi.org/10.1145/335305.335349, https://doi.org/10.1145/335305.335349
10. Janson, S.: Poisson approximation for large deviations. Random Structures and Algorithms **1**(2), 221–229 (1990)
11. Jerrum, M., Snir, M.: Some exact complexity results for straight-line computations over semirings. J. ACM **29**(3), 874–897 (Jul 1982). https://doi.org/10.1145/322326.322341, http://doi.acm.org/10.1145/322326.322341
12. Li, X., Lovett, S., Zhang, J.: Sunflowers and quasi-sunflowers from randomness extractors. In: APPROX-RANDOM. LIPIcs, vol. 116, pp. 51:1–13 (2018)
13. Lovett, S., Solomon, N., Zhang, J.: From dnf compression to sunflower theorems via regularity. arXiv preprint arXiv:1903.00580 (2019)
14. Lovett, S., Zhang, J.: Dnf sparsification beyond sunflowers. In: Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing. pp. 454–460. ACM (2019)
15. Pitassi, T., Robere, R.: Strongly exponential lower bounds for monotone computation. In: Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. pp. 1246–1255. ACM (2017)
16. Rao, A.: Coding for sunflowers (2019), arXiv preprint arXiv:1909.04774
17. Raz, R., Yehudayoff, A.: Multilinear formulas, maximal-partition discrepancy and mixed-sources extractors. J. Comput. Syst. Sci. **77**(1), 167–190 (2011). https://doi.org/10.1016/j.jcss.2010.06.013, https://doi.org/10.1016/j.jcss.2010.06.013
18. Razborov, A.A.: Lower bounds on the monotone complexity of some Boolean functions. Dokl. Akad. Nauk SSSR **281**(4), 798–801 (1985)
19. Rossman, B.: The monotone complexity of k-clique on random graphs. SIAM J. Comput. **43**(1), 256–279 (2014), https://doi.org/10.1137/110839059
20. Srinivasan, S.: Strongly exponential separation between monotone VP and monotone VNP. CoRR **abs/1903.01630** (2019), http://arxiv.org/abs/1903.01630
21. Stichtenoth, H.: Algebraic function fields and codes, vol. 254. Springer Science & Business Media (2009)
22. Tiekenheinrich, J.: A 4n-lower bound on the mononotone network complexity of a oneoutput boolean function. Information Processing Letters **18**, 201–201 (1984)
23. Yehudayoff, A.: Separating monotone VP and VNP. In: Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019. pp. 425–429 (2019). https://doi.org/10.1145/3313276.3316311, https://doi.org/10.1145/3313276.3316311

A Proof of Theorem 4

We include here most of the proof of Theorem 4, which is implicit in [16].

Proof. In what follows, we suppose B is a large enough universal constant.

The proof is by induction on ℓ. Suppose $\ell = 1$. Then $\mathcal{F}$ is a family of singletons. Therefore, the probability that $\boldsymbol{W} \in \binom{[n]}{m}$ chosen uniformly at random does not contain any set of $\mathcal{F}$ is equal to $\binom{n-|\mathcal{F}|}{m}/\binom{n}{m}$. We get

$$\mathop{\mathbb{P}}_{\boldsymbol{W}}\left[\forall F \in \mathcal{F} : F \not\subseteq \boldsymbol{W}\right] = \frac{\binom{n-|\mathcal{F}|}{m}}{\binom{n}{m}} \leqslant \left(\frac{n-m}{n}\right)^{|\mathcal{F}|} \leqslant (1-p/2)^{|\mathcal{F}|} \leqslant e^{-|\mathcal{F}|p/2} \leqslant \varepsilon.$$

Hence, the family $\mathcal{F}$ is itself a (m, ε)-slice-sunflower.

We now proceed by induction, supposing $\ell \geqslant 2$ and that the claim holds for all k-uniform families such that $k < \ell$.

Let $r := Bx \log x$. For any set $T \subseteq [n]$, define

$$\mathcal{F}_T := \{F \setminus T : F \in \mathcal{F} \text{ such that } T \subseteq F\}.$$

We say that $\mathcal{F}$ is *r-well-spread* if $|\mathcal{F}_T| \leqslant r^{\ell-|T|}$ for every non-empty $T \subseteq [n]$. Observe that, if $\mathcal{F}$ is not r-well-spread, then there exists a set $T \subseteq [n]$ such that $|\mathcal{F}_T| \geqslant r^{\ell-|T|}$. Therefore, by the induction hypothesis, $\mathcal{F}_T$ contains a (m, ε)-slice-sunflower $\mathcal{F}'_T$. Observe that the family $\{U \cup T : U \in \mathcal{F}'_T\} \subseteq \mathcal{F}$ is a (m, ε)-slice-sunflower. Therefore, it suffices to consider the case when $\mathcal{F}$ is r-well-spread.

For convenience, let $S_1, \dots, S_t$ be the sets of $\mathcal{F}$. Define $\chi(S_i, \boldsymbol{W})$ to be $S_j \setminus \boldsymbol{W}$, where $j \in [\ell]$ is chosen to minimize $|S_j \setminus \boldsymbol{W}|$ among all choices with $S_j \subseteq S_i \cup \boldsymbol{W}$. If there are many such choices, let j be the smallest one. Note that, for any set $S \in \mathcal{F}$, we have $\chi(S, \boldsymbol{W}) = \emptyset$ if and only if there exists $F \in \mathcal{F}$ such that $F \subseteq \boldsymbol{W}$.

The following key lemma was proved in [16] with a clever coding argument, inspired by the work of Alweiss, Lovett, Wu and Zhang [3].

Lemma 11 ([16]). *For every non-negative integer s, the following holds. Let $\mathcal{F} \subseteq \binom{[n]}{\ell}$ be a r-well-spread family for some $r > 0$. If $\boldsymbol{S}$ is a uniformly random set of the family, and $\boldsymbol{X} \subseteq [n]$ is a uniformly random set of size $s \cdot 128 \cdot \lceil n/r \rceil$ sampled independently, then*

$$\mathop{\mathbb{E}}_{\boldsymbol{X},\boldsymbol{S}}[|\chi(\boldsymbol{S}, \boldsymbol{X})|] \leqslant \ell \cdot (1 - 1/\log r)^s.$$

We now use Lemma 11 to finish the proof. Let $s = \lceil \log(\ell/\varepsilon) \cdot \log r \rceil$. We have

$$\begin{aligned} s \cdot 128 \cdot \lceil n/r \rceil &< 512 \cdot \log r \cdot \log(\ell/\varepsilon) \cdot n/r \\ &= 512 \cdot n \cdot \frac{\log B + \log x + \log\log x}{Bx \log x} \\ &= 512 \cdot np \cdot \frac{\log B + \log x + \log\log x}{B \log(\ell/\varepsilon) \log x} \\ &\leqslant 512 \cdot np \cdot \frac{\log B}{B} < m, \end{aligned}$$

for B large enough. Therefore, by Lemma 11, we get that

$$\begin{aligned} \mathop{\mathbb{E}}_{\boldsymbol{W},\boldsymbol{S}}[|\chi(\boldsymbol{S}, \boldsymbol{W})|] &\leqslant \mathop{\mathbb{E}}_{\boldsymbol{X},\boldsymbol{S}}[|\chi(\boldsymbol{S}, \boldsymbol{X})|] \\ &\leqslant \ell \cdot (1 - 1/\log r)^s \\ &\leqslant \ell \cdot (1 - 1/\log r)^{\log(\ell/\varepsilon) \cdot \log r} \\ &\leqslant \ell e^{-\log(\ell/\varepsilon)} = \varepsilon. \end{aligned}$$

We can conclude the proof by applying Markov's inequality, as follows:

$$\mathop{\mathbb{P}}_{\boldsymbol{W}}[\forall F \in \mathcal{F} : F \not\subseteq \boldsymbol{W}] = \mathop{\mathbb{P}}_{\boldsymbol{W},\boldsymbol{S}}[|\chi(\boldsymbol{S}, \boldsymbol{W})| > 0] \leqslant \mathop{\mathbb{E}}_{\boldsymbol{W},\boldsymbol{S}}[|\chi(\boldsymbol{S}, \boldsymbol{W})|] \leqslant \varepsilon.$$

B Lower Bound for $\mathsf{Clique}_{k,n}$

Recall that the Boolean function $\mathsf{Clique}_{k,n} : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ receives a graph on n vertices as an input and outputs a 1 if this graph contains a clique on k vertices. In this section, we prove an $n^{\Omega(\delta k)}$ lower bound on the monotone circuit size of $\mathsf{Clique}_{k,n}$ for $k = n^{(1/3)-\delta}$.

As in Section 2, we will follow the approximation method. However, instead of using sunflowers as in [18,1] or robust sunflowers as in [19], we introduce a notion of *robust clique-sunflowers* and employ it to bound the errors of approximation.

B.1 Test distributions

We denote by $\boldsymbol{G}_{n,p}$ the Erdős-Rényi random graph, in which each edge appears independently with probability p. Furthermore, fix any $2 \leqslant k = n^{1/3-\delta}$ where $\delta > 0$ and let $p := n^{-2/(k-1)}$. We observe that the probability that $\boldsymbol{G}_{n,p}$ has a k-clique is bounded away from 1.

Lemma 12. *We have* $\mathbb{P}[\ \boldsymbol{G}_{n,p}$ *contains a* k*-clique* $] \leqslant 3/4$.

Proof. There are $\binom{n}{k} \leqslant (en/k)^k$ potential k-cliques, each present in $\boldsymbol{G}_{n,p}$ with probability $p^{\binom{k}{2}} = n^{-k}$. By a union bound, we have $\mathbb{P}[\ \boldsymbol{G}_{n,p}$ contains a k-clique $] \leqslant (e/k)^k \leqslant (e/3)^3 \leqslant 3/4$.

We now define the positive and negative test distributions. For $A \subseteq [n]$, let K_A be the graph on n vertices with a clique on A and no other edges. Let $\boldsymbol{Y}$ be the uniform random graph chosen from all possible K_A. We call $\boldsymbol{Y}$ the *positive test distribution*. Let also $\boldsymbol{N} := \boldsymbol{G}_{n,p}$. We call $\boldsymbol{N}$ the *negative test distribution*. From Lemma 12, we easily obtain the following corollary.

Corollary 3. *We have* $\mathbb{P}[\mathsf{Clique}_{k,n}(\boldsymbol{Y}) = 1] + \mathbb{P}[\mathsf{Clique}_{k,n}(\boldsymbol{N}) = 0] \geqslant 5/4$.

B.2 Robust clique-sunflowers

Here we introduce the notion of *robust clique-sunflowers*, which is analogous to that of robust sunflowers for "clique-shaped" set systems.

Definition 5. *Let* $\varepsilon, p \in (0,1)$*. Let* $\mathcal{S}$ *be a family of subsets of* $[n]$ *and let* $Y := \bigcap \mathcal{S}$*. The family* $\mathcal{S}$ *is called a* (p,ε)-robust clique-sunflower *if*

$$\mathbb{P}\left[\exists A \in S : K_A \subseteq \boldsymbol{G}_{n,p} \cup K_Y\right] \geqslant 1 - \varepsilon.$$

Equivalently, the family $\mathcal{S}$ *is a robust clique-sunflower if the family* $\{K_A : A \in \mathcal{S}\} \subseteq \binom{[n]}{2}$ *is a* (p,ε)*-robust sunflower, since* $K_A \cap K_B = K_{A\cap B}$.

Though clique-sunflowers may seem similar to regular sunflowers, the importance of this definition is that it allows us to explore the "clique-shaped" structure of the sets of the family, and thus obtain an asymptotically better upper bound on the size of sets that do not contain a robust clique-sunflower.

Lemma 13. *Let* $\mathcal{S}$ *be such that* $|\mathcal{S}| \geqslant \ell!(2\ln(1/\varepsilon))^\ell (1/p)^{\binom{\ell}{2}}$*. Then* $\mathcal{S}$ *contains a* (p,ε)*-robust clique-sunflower.*

Observe that, whereas the bounds for "standard" robust sunflowers (Theorems 2, 3, 4) would give us an exponent of $\binom{\ell}{2}$ on the $\log(1/\varepsilon)$ factor, Lemma 13 give us only an ℓ at the exponent. As we shall see, this is asymptotically better for our choice of parameters.

We defer the proof of Lemma 13 to the next section. The proof is based on an application of Janson's inequality [10], as in the original robust sunflower lemma of [19] (Theorem 2). We expect that a proof along the lines of the work of Alweiss *et al* [3] and Rao [16] should be able to give us an even better bound, removing the $\ell!$ factor. This would extend our $n^{\Omega(k)}$ lower bound to $k \leqslant n^{1/2-o(1)}$.

B.3 A closure operator

As in Section 2.3, we define here a closure operator in the lattice of monotone Boolean functions. We will again prove that the closure of a function will be a good approximation for it on the negative test distribution. However, unlike Section 2.3, instead of bounding the set of minterms, we will bound the set of "clique-shaped" minterms, as we shall see. Throughout this section, we fix the error parameter

$$\varepsilon := n^{-k}.$$

Definition 6. *We say that* $f \in \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ *is* ε-closed *if, for every* $A \in \binom{[n]}{\leqslant k}$*, we have*

$$\mathbb{P}[\ f(\boldsymbol{N} \cup K_A) = 1\ ] \geqslant 1 - \varepsilon \implies f(K_A) = 1.$$

As before, we can define the closure $\mathrm{cl}(f)$ of a monotone Boolean function f, and bound the error of approximating f by $\mathrm{cl}(f)$ under $\boldsymbol{N}$.

Lemma 14. *For every monotone* $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$*, we have*

$$\mathbb{P}\left[f(\boldsymbol{N}) = 0 \text{ and } \mathrm{cl}(f)(\boldsymbol{N}) = 1\right] \leqslant \varepsilon \sum_{j=0}^{\delta k} \binom{n}{j} \leqslant \varepsilon n^{\delta k} \leqslant n^{-(2/3)k}.$$

Proof. Same as the proof of Lemma 4.

Let $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ be monotone and suppose $\ell \in [k]$. We define

$$\mathcal{M}_\ell(f) := \{A \in \tbinom{[n]}{\ell} : f(K_A) = 1 \text{ and } f(K_{A\setminus\{a\}}) = 0 \text{ for all } a \in A\}.$$

Elements of $\mathcal{M}_\ell(f)$ are called *ℓ-clique-minterms* of f. By employing the robust clique-sunflower lemma (Lemma 13), we are able to bound the set of ℓ-clique-minterms of closed monotone functions.

Lemma 15. *If a monotone function* $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ *is ε-closed, then, for all* $\ell \in [k]$*, we have*

$$|\mathcal{M}_\ell(f)| \leqslant \frac{(2\ell \log(1/\varepsilon))^\ell}{p^{\binom{\ell}{2}}}.$$

Proof. Same as the proof of Lemma 5.

B.4 Trimmed monotone functions

In this section, we define again a trimming operation for Boolean functions and prove analogous bounds to that of Section 2.4.

Definition 7. *We say that a function* $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ *is* clique-shaped *if, for every minterm x of f, there exists $A \subseteq [n]$ such that $x = K_A$ (that is, every minterm of f is a clique). Moreover, we say that f is* trimmed *if f is clique-shaped and all the clique-minterms of f have size at most $\delta k/2$.*

For a set $A \in \binom{[n]}{\leqslant k}$, let $\lceil A \rceil : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ denote the indicator function of containing K_A, which satisfies

$$\lceil A \rceil(G) = 1 \iff K_A \subseteq G.$$

Functions of the forms $\lceil A \rceil$ are called *clique-indicators*. Moreover, if $|A| \leqslant \ell$, we say that $\lceil A \rceil$ is a clique-indicator of size at most ℓ. Let $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ be clique-shaped. We define

$$\mathrm{trim}(f) := \bigvee_{\ell=2}^{\delta k/2} \bigvee_{A \in \mathcal{M}_\ell(f)} \lceil A \rceil.$$

That is, the trim operation takes out from f all the clique-indicators of size larger than δk, yielding a trimmed function.

Note that the probability that a random K_A sampled from $\boldsymbol{Y}$ contains one of the clique-minterms of size ℓ of a function f is at most

$$\frac{\binom{n-k}{k-l}}{\binom{n}{k}} |\mathcal{M}_\ell(f)| \leqslant \left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(f)|.$$

Imitating the proofs of Lemmas 6 and 7, we may now obtain the following lemmas.

Lemma 16. *If a monotone function* $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ *is a trimmed clique-shaped function such that* $f \neq \mathbb{1}$, *then*

$$\mathbb{P}\left[f(\boldsymbol{Y}) = 1\right] \leqslant \sum_{\ell=2}^{\delta k/2} \left(\frac{k}{n}\right)^{\ell} |\mathcal{M}_\ell(f)|.$$

Lemma 17. *Let* $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ *be a clique-shaped monotone function, all of whose clique-minterms have size at most* δk. *We have*

$$\mathbb{P}\left[f(\boldsymbol{Y}) = 1 \text{ and } \operatorname{trim}(f)(\boldsymbol{Y}) = 0\right] \leqslant \sum_{\ell=\delta k/2}^{\delta k} \left(\frac{k}{n}\right)^{\ell} |\mathcal{M}_\ell(f)|.$$

B.5 Approximators

Similarly as in the previous lower bound, we will consider a set of *approximators* $\mathcal{A}$. Let $\mathcal{A} := \{\operatorname{trim}(\operatorname{cl}(f)) : f \in \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ is monotone and clique-shaped$\}$. We define operations $\sqcup, \sqcap : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ as follows: for $f, g \in \mathcal{A}$ such that $f = \bigvee_{i=1}^{t} \lceil A_i \rceil$ and $g = \bigvee_{i=j}^{s} \lceil B_j \rceil$, let

$$\begin{aligned} f \sqcup g &:= \operatorname{trim}(\operatorname{cl}(f \vee g)), \\ f \sqcap g &:= \operatorname{trim}(\operatorname{cl}(\bigvee_{i,j} \lceil A_i \cup B_j \rceil)). \end{aligned}$$

For convenience, we let $\bigwedge(f,g) := \bigvee_{i,j} \lceil A_i \cup B_j \rceil$. Observe that every edge-indicator $\lceil \{u,v\} \rceil$ belongs to $\mathcal{A}$. If C is a monotone $\{\vee, \wedge\}$-circuit, let $C^{\mathcal{A}}$ be the corresponding $\{\sqcup, \sqcap\}$-circuit, which computes an approximator.

B.6 The lower bound

In this section we finally obtain the desired lower bound for the clique function. We will prove that, if $k \leqslant n^{1/3-\delta}$ for some constant $\delta > 0$, then the monotone complexity of $\mathsf{Clique}_{k,n}$ is $n^{\Omega(k)}$. Henceforth, we will suppose that this is the case. We begin by defining, for every $2 \leqslant \ell \leqslant \delta k$, the number

$$s_\ell := \left(\frac{k}{n}\right)^{\ell} \frac{(2\ell \log(1/\varepsilon))^{\ell}}{p^{\binom{\ell}{2}}} = \frac{(2\ell k^2 \log n)^{\ell}}{n^{\ell} p^{\binom{\ell}{2}}}.$$

By Lemma 15, we get that, for every ε-closed monotone function $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$, we have

$$\left(\frac{k}{n}\right)^{\ell} |\mathcal{M}_\ell(f)| \leqslant s_\ell.$$

As seen in Section 2.6, it will be important for us to upper bound the values of s_2 and $s_{\ell+1}/s_\ell$ for all $2 \leqslant \ell < \delta k$, which we now do:

$$\begin{aligned} s_2 &= \left(\frac{k}{n}\right)^2 \frac{(12k \log n)^2}{p} = \left(\frac{O(k^2 \log n)}{n^{1-(1/(k-1))}}\right)^2 \leqslant \left(\frac{O(\log n)}{n^{(1/3)+2\delta}}\right)^2 = o\left(\frac{1}{n^{1/2}}\right), \\ s_{\ell+1}/s_\ell &= \frac{6k^2 \log n}{n} \cdot \frac{(\ell+1)^{\ell+1}}{(\ell p)^{\ell}} \leqslant \frac{O(k^3 \log n)}{np^{\ell}} \leqslant \frac{O(\log n)}{n^{(1/3)+3\delta-(2\ell/(k-1))}} \leqslant \frac{O(\log n)}{n^{(1/3)+\delta}} \leqslant o\left(\frac{1}{n^{1/4}}\right). \end{aligned}$$

It follows that $s_\ell \leqslant O(n^{-\ell/4})$ for all $2 \leqslant \ell \leqslant \delta k$.

Repeating the same arguments of Lemmas 8 and 9, we obtain the following analogous lemmas.

Lemma 18 (Approximators make many errors). *For every $f \in \mathcal{A}$, we have*

$$\mathbb{P}[f(\boldsymbol{Y}) = 1] + \mathbb{P}[f(\boldsymbol{N}) = 0] \leqslant 1 + o(1).$$

Proof. Let $f \in \mathcal{A}$. By definition, there exists an ε-closed function h such that $f = \mathrm{trim}(h)$. Observe that $\mathcal{M}_\ell(f) \subseteq \mathcal{M}_\ell(h)$ for every $\ell \in [c]$. By Lemma 16, if $f \in \mathcal{A}$ such that $f \neq \mathbb{1}$, then

$$\mathbb{P}[f(\boldsymbol{Y}) = 1] \leqslant \sum_{\ell=1}^{\delta k/2} \left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(h)| \leqslant \sum_{\ell=1}^{\delta k/2} s_\ell \leqslant \sum_{\ell=1}^{\delta k/2} O(n^{-\ell/4}) \leqslant o(1).$$

Therefore, for every $f \in \mathcal{A}$ we have $\mathbb{P}[f(\boldsymbol{Y}) = 1] + \mathbb{P}[f(\boldsymbol{N}) = 0] \leqslant 1 + o(1)$.

Lemma 19 (C is well-approximated by $C^{\mathcal{A}}$). *Let C be a monotone circuit. We have*

$$\mathbb{P}[C(\boldsymbol{Y}) = 1 \textit{ and } C^{\mathcal{A}}(\boldsymbol{Y}) = 0] + \mathbb{P}[C(\boldsymbol{N}) = 0 \textit{ and } C^{\mathcal{A}}(\boldsymbol{N}) = 1] \leqslant \mathrm{size}(C) \cdot O(n^{-\delta k/8}).$$

Proof. To bound the approximation errors under the distribution $\boldsymbol{Y}$, first note that, if $f, g \in \mathcal{A}$, then all the clique-minterms of both $f \vee g$ and $f \wedge g$ have Hamming weight at most δk. Moreover, if $(f \vee g)(x) = 1$ but $(f \sqcup g)(x) = 0$, then $\mathrm{trim}(\mathrm{cl}(f \vee g)(x)) \neq \mathrm{cl}(f \vee g)(x)$. Therefore, we obtain by Lemma 17 that, for $f, g \in \mathcal{A}$, we have

$$\mathbb{P}\left[(f \vee g)(\boldsymbol{Y}) = 1 \text{ and } (f \sqcup g)(\boldsymbol{Y}) = 0\right] \leqslant \sum_{\ell=\delta k/2}^{\delta k} \left(\frac{k}{n}\right)^\ell |\mathcal{M}_\ell(f \vee g)| \leqslant \sum_{\ell=\delta k/2}^{\delta k} s_\ell \leqslant \sum_{\ell=\delta k/2}^{\delta k} O(n^{-\ell/4}) = O(n^{-\delta k/8}).$$

The same argument shows $\mathbb{P}\left[(f \wedge g)(\boldsymbol{Y}) = 1 \text{ and } (f \sqcap g)(\boldsymbol{Y}) = 0\right] = O(n^{-\delta k/8})$, which implies

$$\mathbb{P}[C(\boldsymbol{Y}) = 1 \text{ and } C^{\mathcal{A}}(\boldsymbol{Y}) = 0] \leqslant \mathrm{size}(C) \cdot O(n^{-\delta k/8}).$$

Similarly, to bound the approximation errors under $\boldsymbol{N}$, note that $(f \vee g)(x) = 0$ and $(f \sqcup g)(x) = 1$ only if $\mathrm{cl}(f \vee g)(x) \neq (f \vee g)(x)$. Therefore, we obtain by Lemma 14 that, for $f, g \in \mathcal{A}$, we have

$$\mathbb{P}\left[(f \vee g)(\boldsymbol{N}) = 0 \text{ and } (f \sqcup g)(\boldsymbol{N}) = 1\right] \leqslant n^{-(2/3)k}.$$

By the same argument above, we obtain

$$\mathbb{P}[C(\boldsymbol{N}) = 0 \text{ and } C^{\mathcal{A}}(\boldsymbol{N}) = 1] \leqslant \mathrm{size}(C) \cdot n^{-(2/3)k} \leqslant \mathrm{size}(C).$$

This finishes the proof.

We can finally obtain the lower bound for the clique function.

Theorem 9. *For all $k = n^{1/3-\delta}$ where $0 < \delta < 1/3$, the monotone circuit complexity of $\mathsf{Clique}_{k,n}$ is $\Omega(n^{\delta k/8})$.*

Proof. Let C be a monotone circuit computing $\mathsf{Clique}_{k,n}$. We have

$$\begin{aligned}
5/4 &\leqslant \mathbb{P}[\mathsf{Clique}_{k,n}(\boldsymbol{Y})] + \mathbb{P}[\mathsf{Clique}_{k,n}(\boldsymbol{N})] \\
&\leqslant \mathbb{P}[C(\boldsymbol{Y}) = 1 \text{ and } C^{\mathcal{A}}(\boldsymbol{Y}) = 0] + \mathbb{P}[C^{\mathcal{A}}(\boldsymbol{Y}) = 1] \\
&\quad + \mathbb{P}[C(\boldsymbol{N}) = 0 \text{ and } C^{\mathcal{A}}(\boldsymbol{N}) = 1] + \mathbb{P}[C^{\mathcal{A}}(\boldsymbol{N}) = 1] \\
&\leqslant 1 + o(1) + \mathrm{size}(C) \cdot O(n^{-\delta k/8}).
\end{aligned}$$

This implies $\mathrm{size}(C) = \Omega(n^{\delta k/8})$.

C Proof of Lemma 13

Let $\boldsymbol{U}_{n,q} \subseteq [n]$ be a q-random subset of $[n]$ (independent of $\boldsymbol{G}_{n,p}$). Let $c_1 := \ln(1/\varepsilon)$ and for $\ell \geqslant 2$, let $c_\ell := 2\ln(1/\varepsilon)\sum_{j=1}^{\ell-1}\binom{\ell}{j}c_j$. The following can be easily checked.

Lemma 20. $c_\ell \leqslant \ell!(2\log(1/\varepsilon))^\ell$.

It follows from the definition of robust clique-sunflowers that the robust clique-sunflower lemma (Lemma 13) is implied by the following result.

Lemma 21. *For all* $\ell \in \{1,\ldots,n\}$ *and* $S \subseteq \binom{[n]}{\ell}$, *if* $|S| \geqslant c_\ell(1/q)^\ell(1/p)^{\binom{\ell}{2}}$, *then there exists* $B \in \binom{[n]}{<\ell}$ *such that*

$$\mathbb{P}[\bigwedge_{A\in S\,:\,B\subseteq A}(K_A \nsubseteq \boldsymbol{G}_{n,p}\cup K_B \text{ or } A \nsubseteq \boldsymbol{U}_{n,q}\cup B)\ ] \leqslant \varepsilon.$$

Proof. By induction on ℓ. In the base case $\ell = 1$, we have $B = \emptyset$ and (by independence)

$$\begin{aligned}\mathbb{P}[\bigwedge_{A\in S}(K_A \nsubseteq \boldsymbol{G}_{n,p} \text{ or } A \nsubseteq \boldsymbol{U}_{n,q})\ ] &= \mathbb{P}[\bigwedge_{A\in S}(A \nsubseteq \boldsymbol{U}_{n,q})\ ]\\ &= \prod_{A\in S}\mathbb{P}[\ A \nsubseteq \boldsymbol{U}_{n,q}\ ]\\ &= (1-q)^{|S|} \leqslant (1-q)^{\ln(1/\varepsilon)/q} \leqslant e^{-\ln(1/\varepsilon)} = \varepsilon.\end{aligned}$$

Let $\ell \geqslant 2$. First, consider the case that there exists $j \in \{1,\ldots,\ell-1\}$ and $B \in \binom{[n]}{j}$ such that

$$|\{A \in S : B \subseteq A\}| \geqslant c_{\ell-j}(1/qp^j)^{\ell-j}(1/p)^{\binom{\ell-j}{2}}.$$

Let $T = \{A \setminus B : A \in S \text{ such that } B \subseteq A\} \subseteq \binom{[n]}{\ell-j}$. By the induction hypothesis, there exists $D \in \binom{[n]\setminus B}{<\ell-j}$ such that

$$\mathbb{P}[\bigwedge_{C\in T\,:\,D\subseteq C}(K_C \nsubseteq \boldsymbol{G}_{n,p}\cup K_D \text{ or } C \nsubseteq \boldsymbol{U}_{n,qp^j}\cup D)\ ] \leqslant \varepsilon.$$

We have

$$\begin{aligned}&\mathbb{P}[\bigwedge_{A\in S\,:\,B\cup D\subseteq A}(K_A \nsubseteq \boldsymbol{G}_{n,p}\cup K_{B\cup D} \text{ or } A \nsubseteq \boldsymbol{U}_{n,q}\cup B\cup D)\ ]\\ &= \mathbb{P}[\bigwedge_{C\in T\,:\,D\subseteq C}(K_{B\cup C} \nsubseteq \boldsymbol{G}_{n,p}\cup K_{B\cup D} \text{ or } B\cup C \nsubseteq \boldsymbol{U}_{n,q}\cup B\cup D)\ ]\\ &= \mathbb{P}[\bigwedge_{C\in T\,:\,D\subseteq C}(K_{B\cup C} \nsubseteq \boldsymbol{G}_{n,p}\cup K_{B\cup D} \text{ or } C \nsubseteq \boldsymbol{U}_{n,q}\cup D)\ ]\\ &= \mathbb{P}[\bigwedge_{C\in T\,:\,D\subseteq C}(K_C \nsubseteq \boldsymbol{G}_{n,p}\cup K_D \text{ or } C \nsubseteq \big\{v\in \boldsymbol{U}_{n,q} : \{v,w\}\in E(\boldsymbol{G}_{n,p}) \text{ for all } w\in B\big\}\cup D\ ]\\ &\leqslant \mathbb{P}[\bigwedge_{C\in T\,:\,D\subseteq C}(K_C \nsubseteq \boldsymbol{G}_{n,p}\cup K_D \text{ or } C \nsubseteq \boldsymbol{U}_{n,qp^j}\cup D\ ]\\ &\leqslant \varepsilon.\end{aligned}$$

Finally, assume that for all $j \in \{1,\ldots,\ell-1\}$ and $B \in \binom{[n]}{j}$, we have

$$|\{A \in S : B \subseteq A\}| \leqslant c_{\ell-j}(1/qp^j)^{\ell-j}(1/p)^{\binom{\ell-j}{2}}.$$

In this case, we show that the bound of the lemma holds with $B = \emptyset$. Let

$$\mu := |S| q^{\ell} p^{\binom{\ell}{2}},$$
$$\Delta := \sum_{j=1}^{\ell-1} \sum_{(A,A') \in S^2 \,:\, |A \cap A'| = j} q^{2\ell - j} p^{2\binom{\ell}{2} - \binom{j}{2}}.$$

Janson's Inequality [10] gives the following bound:

$$\mathbb{P}[\bigwedge_{A \in S} (K_A \nsubseteq \boldsymbol{G}_{n,p} \text{ or } A \nsubseteq \boldsymbol{U}_{n,q})\ ] \leqslant \exp\left(-\frac{\mu^2}{\mu + \Delta}\right). \tag{1}$$

We bound Δ as follows:

$$\begin{aligned}
\Delta &\leqslant \sum_{j=1}^{\ell-1} q^{2\ell-j} p^{2\binom{\ell}{2} - \binom{j}{2}} \sum_{B \in \binom{[n]}{j}} |\{A \in S : B \subseteq A\}|^2 \\
&\leqslant \sum_{j=1}^{\ell-1} q^{2\ell-j} p^{2\binom{\ell}{2} - \binom{j}{2}} \sum_{B \in \binom{[n]}{j}} |\{A \in S : B \subseteq A\}| \cdot c_{\ell-j} (1/q)^{\ell-j} (1/p)^{\binom{\ell-j}{2}} \\
&= q^{\ell} p^{\binom{\ell}{2}} \sum_{j=1}^{\ell-1} c_{\ell-j} \sum_{B \in \binom{[n]}{j}} |\{A \in S : B \subseteq A\}| \\
&= q^{\ell} p^{\binom{\ell}{2}} \sum_{j=1}^{\ell-1} c_{\ell-j} \sum_{A \in S} \sum_{B \in \binom{A}{j}} 1 \\
&= |S| q^{\ell} p^{\binom{\ell}{2}} \sum_{j=1}^{\ell-1} \binom{\ell}{j} c_{\ell-j} \\
&= \mu \sum_{j=1}^{\ell-1} \binom{\ell}{j} c_j.
\end{aligned}$$

We next observe that $\mu \geqslant c_\ell$, since $|S| \geqslant c_\ell (1/q)^{\ell} (1/p)^{\binom{\ell}{2}}$. Therefore,

$$\frac{\mu^2}{2\Delta} \geqslant \frac{\mu}{2 \sum_{j=1}^{\ell-1} \binom{\ell}{j} c_{\ell-j}} = \frac{|S| q^{\ell} p^{\binom{\ell}{2}}}{2 \sum_{j=1}^{\ell-1} \binom{\ell}{j} c_{\ell-j}} \geqslant \frac{c_\ell}{2 \sum_{j=1}^{\ell-1} \binom{\ell}{j} c_{\ell-j}} = \ln(1/\varepsilon).$$

Finally, from (1) we get

$$\mathbb{P}[\bigwedge_{A \in S} (K_A \nsubseteq \boldsymbol{G}_{n,p} \text{ or } A \nsubseteq \boldsymbol{U}_{n,q})\ ] \leqslant \exp\left(-\frac{\mu^2}{2\Delta}\right) \leqslant \varepsilon.$$